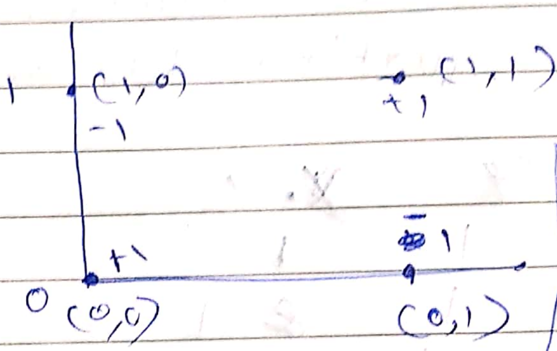
Q3) XOR

A	B	Output
1	1	1
0	1	0
1	0	0
0	0	1



A	B	Output
1	1	1
-1	1	-1
1	-1	-1
-1	-1	1

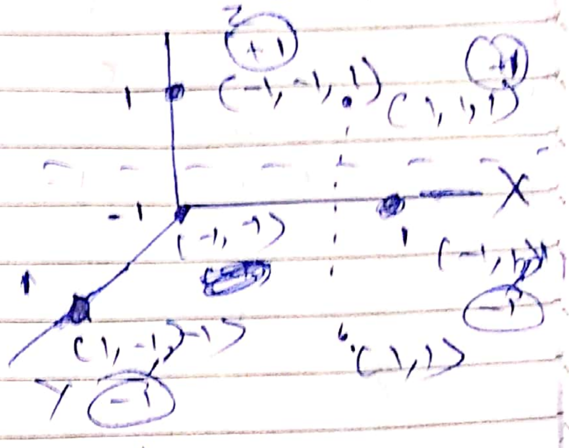
Assuming 0 = -1 class.

Clearly acc to figure classes are not linearly separable.

=> A kernel is needed.

If we use a new attribute $Z = X \cdot Y$ then the equation becomes linearly separable.

X	Y	X.Y
1	1	1
-1	1	-1
1	-1	-1
-1	-1	1



$\Rightarrow \phi$ Map: (x, y)
$\rightarrow (x, y, xy)$

$\Rightarrow$ Kernel $K(x_i, x_j) = [x_i \; y_i \; x_i y_i] \begin{bmatrix} x_j \\ y_j \\ x_j y_j \end{bmatrix}$

$= x_i x_j + y_i y_j + x_i x_j y_i y_j$

Gradhan Matrix

$$\begin{array}{c|cccc} & (1,1) & (-1,1) & (1,-1) & (-1,-1) \\ \hline (1,1) & 3 & -1 & -1 & -1 \\ K = (-1,1) & -1 & 3 & -1 & -1 \\ (1,-1) & -1 & -1 & 3 & -1 \\ (-1,-1) & -1 & -1 & -1 & 3 \end{array}$$

$1 = (1, 1)$
$2 = (-1, 1)$
$3 = (1, -1)$
$4 = (-1, -1)$

$$L_{ds}(\alpha) = \sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i=1}^{4} \sum_{j=1}^{4} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2} \Big(3\alpha_1^2 + \alpha_1\alpha_2 + \alpha_1\alpha_3 - \alpha_1\alpha_4 + 3\alpha_2^2 + \alpha_1\alpha_2 + \alpha_2\alpha_4 - \alpha_2\alpha_3 + 3\alpha_3^2 + \alpha_3\alpha_1 + \alpha_3\alpha_4 - \alpha_3\alpha_2 + 3\alpha_4^2 + \alpha_4\alpha_2 + \alpha_4\alpha_3 - \alpha_4\alpha_1 \Big)$$

$\Rightarrow$ ~~$\alpha_1(3\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4) =$~~

$\Rightarrow$ On diff $\text{Loss}(\alpha) = \frac{d}{d\alpha}\left(\alpha - \frac{1}{2}H(\alpha^T\alpha)\right)$

where H = Gram matrix $\times (y_i y_j)$

$\Rightarrow 1 - H\alpha = 0$

$\Rightarrow H\alpha = 1$

$\Rightarrow 3\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 1$

$\alpha_1 + 3\alpha_2 + \alpha_3 + \alpha_4 = 1$

$\alpha_1 + \alpha_2 + 3\alpha_3 + \alpha_4 = 1$

$-\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 = 1$

$\Rightarrow$ optimal $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{4}$

$\Rightarrow$ All are support vectors,
Sub values in Loss equation we get

$$\text{Loss}(\alpha) = 1 - \frac{1}{2}\left(\frac{3}{16}\times 4 + \frac{1}{16}\times 4\right)$$

$$= 1 - \frac{1}{2}\left(4 \times \frac{4}{16}\right)$$

$$= 1 - \frac{1}{2} = \frac{1}{2}$$

$\Rightarrow \vec{W} = \sum_{i=1}^{4} \alpha_i y_i \phi(x_i)$

$$= \frac{1}{4}\big((1,1,1) + (-1,1,-1) + (1,-1,-1) + (-1,-1,1)\big)$$

$$= \frac{1}{4}\big((0,0,4)\big) = (0,0,1)$$

$\Rightarrow$ boundary $= \vec{w}\cdot\phi(x) = (0,0,1)\cdot(x, y, x\cdot y)$

$= x\cdot y = 0$

if $x = y = 1$ and $x = y = -1$
$\Rightarrow x.y = 1 \times 1 > 0$ & $-1 \times -1 > 0$
$\Rightarrow$ class $= 1$
if $x = 1, y = -1$ & $x = -1, y = 1$
$\Rightarrow x.y = -1 \times 1 < 0$ & $1 \times -1 < 0$
done $\Rightarrow$ class $= -1$
hence XOR done by SVM

Q2) $k(x_i, x_j) = (1 + x^T . x')^2$

where $x^T = [x_1 \; x_2]$
& $x' = \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$

$\Rightarrow x^T . x' = [x_1 \; x_2] \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = x_1 x'_1 + x_2 x'_2$

$\Rightarrow (1 + x_1 x'_1 + x_2 x'_2)^2$

$= (1.1) + (x_1^2)(x_1'^2) + (x_2^2)(x_2'^2) + 2(x_1)(x_1') + 2(x_2)(x_2')$
$+ 2(x_1 x_1')(x_2 x_2')$

$\phi(x) . \phi(x') = k(x_i, x_j)$

As we have 6 terms in the expansion of kernel function & each term can be expressed as the product of 2 terms.

$1 = 1 . 1$, $x_1^2 x_1'^2 = x_1^2 \times x_1'^2$,
$(x_2^2 . x_2'^2) = x_2^2 \times x_2'^2$
$2 x_1 x_1' = \sqrt{2} x_1 \times \sqrt{2} x_1'$

$2x_2x_2' = \sqrt{2}x_2 \times \sqrt{2}x_2'$
$2x_1x_1'x_2x_2' = \sqrt{2}x_1x_2 \times \sqrt{2}x_1'x_2'$

where each term is a product of one term in x & one term in x'.

$$\Rightarrow k(x, x') = [1, x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2]^T \cdot [1, x_1'^2, x_2'^2, \sqrt{2}x_1', \sqrt{2}x_2', \sqrt{2}x_1'x_2']$$

$$\Rightarrow \phi(x) = [1, x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2]^T$$

There are 6 features in $\phi(x)$

Q3) $$w = \sum_{i=1}^{m} \alpha_i y_i x_i = \sum_{i=1}^{10} \alpha_i y_i x_i$$

$$= 0.414 \times 1 \times (4, 2.9) + 0 + 0 + (0.18 \times -1 \times (2.5, 1)) + 0 + 0 + (0.018 \times 1 \times (3.5, 4)) + 0 + (0.414 \times -1 \times (2, 2.1)) + 0$$

$$= \begin{bmatrix} 1.656 \\ 1.2006 \end{bmatrix} + \begin{bmatrix} -2.45 \\ -1.18 \end{bmatrix} + \begin{bmatrix} 0.063 \\ 0.072 \end{bmatrix} + \begin{bmatrix} -0.828 \\ -0.8694 \end{bmatrix} = \begin{bmatrix} -2.059 \\ -0.7768 \end{bmatrix}$$

2) Support vectors $\Rightarrow \alpha_i \neq 0$
$\Rightarrow x_i = (4, 2.9), (2.5, 1), (3.5, 4), (2, 2.1)$ are all support vectors.

Using KKT Condition

$$b = 1 - W^T x^{(s)}$$

$b_1 = 1 - W^T \begin{bmatrix} 4 \\ 2.9 \end{bmatrix} = 11.48872$

$b_2 = 1 - W^T \begin{bmatrix} 2.5 \\ 1 \end{bmatrix} = 1 - (-5.9263) = 6.9263$

$b_3 = 1 - W^T \begin{bmatrix} 3.5 \\ 4 \end{bmatrix} = 11.3137$

$b_4 = 1 - W^T \begin{bmatrix} 2 \\ 2.1 \end{bmatrix} = 1 - (-5.74928) = 6.74928$

$\Rightarrow$ avg basis $= 9.119$

for $(3,3)$

$\Rightarrow \quad W^T x + b \geq 0$

$\Rightarrow -8.5074 + 9.119 = 0.6116$

$\Rightarrow -1$ class

Q6) **BONUS**

RBF kernel $= e^{-\gamma\|u-v\|^2}$

$\|u-v\| = \sqrt{(u_1-v_1)^2+(u_2-v_2)^2} \dots$

but u & v are 1-D

$\Rightarrow \|u-v\| = \sqrt{(u-v)^2} = (u-v)$

$\Rightarrow$ Kernal $= e^{-\gamma(u-v)^2} = \phi(u)\,\phi(v)$

$= e^{-\gamma(u^2+v^2-2uv)}$

$= e^{\gamma(-u^2-v^2+2uv)}$

$= e^{\gamma(-u^2-v^2)} \times e^{\gamma(2uv)}$

using Taylor expansion of $e^{\gamma(2uv)}$

$= \left(1+\frac{\gamma 2uv}{1!}+\frac{(\gamma 2uv)^2}{2!}+\dots\right)$

$= 1\cdot 1+\frac{\sqrt{2\gamma}u\cdot\sqrt{2\gamma}v}{1!\;1!}+\sqrt{\frac{(2\gamma)^2u^2}{2!}}\times\frac{\sqrt{2\gamma}\cdot v^2}{\sqrt{2!}} + \dots)$

where each term is a dot product of terms one from u & one from v.

$\Rightarrow K(u,v) = e^{-u^2\gamma}\times e^{-v^2\gamma}\times\left(1\cdot1+\frac{\sqrt{2\gamma}u}{1!}\cdot\frac{\sqrt{2\gamma}v}{1!}+\dots\right)$

$= e^{-u^2\gamma}\left(1,\frac{\sqrt{2\gamma}u}{1!},\sqrt{\frac{(2\gamma)^2u^2}{2!}}\dots\right)\cdot\left(e^{-v^2\gamma}\left(1,\frac{\sqrt{2\gamma}v}{1!}\dots\right)\right)$

Date
Page

$\Rightarrow \phi(\cancel{u}) = e^{-\gamma u^2}\left(1, \sqrt{\frac{2\gamma}{1!}}u, \sqrt{\frac{(2\gamma)^2}{2!}}u^2 \ldots\right)$

$\phi(v)$

$\phi(v) = e^{-\gamma v^2}\left(1, \sqrt{\frac{2\gamma}{1!}}v, \sqrt{\frac{(2\gamma)^2}{2!}}v^2 \ldots\right)$

$\Rightarrow K(u, v) = \phi(u) . \phi(v)$

$\Rightarrow$ When transform

$\phi(x) = e^{-\gamma x^2}\left(1, \sqrt{\frac{2\gamma}{1!}}x, \sqrt{\frac{(2\gamma)^2}{2!}}x^2, \ldots\right)$

as there is a factorial in the denominator, when the ~~deg~~ degree of the terms $\simeq$ ~~100000~~ 1000 000, then the denominator is going to be $\simeq \infty$, but the numerator is although quite big, it is not as big as the denominator. $\Rightarrow$ the terms goes to zero.

Hence higher degrees are close to zero (0) in rbf kernel.